

Experimental evidence for 2SG direct object pronoun choice in Brazilian Portuguese

Scott A. Schwenter, Justin Bland, Kendra Dickinson, Mark Hoff, and Luana Lamberti
The Ohio State University

Few grammars acknowledge the variability between clitic second-person singular (2SG) *te* and tonic pronoun *você* as direct object (DO) pronouns in Brazilian Portuguese (BP), as in *Eu te vi* ~ *Eu vi você*. Those that do mention it offer no explanation for the factors conditioning the variation. Perini (2010), for instance, says that the two pronouns vary “livremente” as DOs. However, Zampaulo’s (2014) corpus study identifies significant differences in *te/você* use by VP complexity, with more *você* in complex VPs (those with 2 or 3 verbs). He also mentions the possible import of contrast but provides no empirical evidence for its effects. In a perception study, Lamberti et al. (2017) found that while 1st and 3rd-person tonic subject pronouns for DO function are stigmatized in BP, the use of tonic *você* is much less so. Our experimental study builds on this prior research through a controlled investigation of the variability in 2SG DO pronoun expression which demonstrates that subject pronoun preference, contrast, and negation all play a significant role in conditioning this choice.

Data were collected via an online forced-choice survey on Qualtrics where respondents chose between *te* and *você* as DO pronoun in written sentential stimuli. A total of 146 native BP survey respondents answered demographic questions and provided their preferred subject pronoun (*você/tu*) before completing 30 survey items (15 target, 15 filler). These data were submitted to mixed-effects logistic regression analyses in R, with participant and item as random effects. We considered the following internal variables: polarity, two types of contrast (single: DO only; double: both DO and predicate), VP complexity, clitic position (+/- sentence initial), and participant-reported 2SG subject pronoun preference (*você* vs. *tu*). We also included the external factors respondent age, gender, socioeconomic class, level of education, and state of origin.

Our data (N=2119) show that *te* (67.7%; n=1434) is selected as a DO more often than *você* (32.3%, n=685) overall. However, this preference is much stronger among respondents who report using *tu* as their preferred subject pronoun (93.3% *te*; n=251) than for those who use *você* (64% *te*; n=1183). Furthermore, results of five different best-fitting mixed-effects logistic regression model reveal additional linguistic constraints on DO pronoun choice. First, we found a statistically significant difference in DO choice between contrastive and non-contrastive contexts ($p < 2e-16$). In non-contrastive contexts, *te* is the default selection (75.7% *te*; n=621), but in contrastive contexts, preference for *você* is much higher (52.1%; n=294). Second, we find *te* is statistically significantly less likely to be selected ($p=0.00772$) when doing so would result in sentence-initial *te*, which is prescriptively prohibited. In such cases, *você* is selected more often (33.5% *você*; n=48) than when choice of *te* would not result in a sentence-initial clitic (22.1% *você*; n=31). Finally, our results show that respondents choose *você* significantly more often ($p=8.50e-05$) in sentences with double contrast (74.1% *você*; n=100) than in sentences with single contrast (50% *você*; n=71). Number of verbs, specific negative polarity item, and were statistically significant constraints on DO choice in these data. However, none of the social factors considered were significant predictors.

Our findings elucidate a set of clear probabilistic constraints on *te* and *você* as DO pronouns; they are not used “livremente” as grammars would have it. While *te* is preferred overall, *você* is the variant preferred in contrastive contexts. Moreover, the stronger the contrast, the more likely *você* is to be chosen by respondents. An important theoretical consequence of this finding

is that it reveals pronominal contrast not to be a binary category, but rather a gradient one: the prosodic prominence that tonic *você*, but not clitic *te*, can carry becomes more crucial the greater the contrast involved. Paradigmatic pressures may also be at work, given that tonic pronouns in spoken BP have already replaced clitics in near-categorical fashion for 3SG DOs (Schwenter 2014), and that in the plural tonic *vocês* is the only option for DO pronoun expression in spoken BP. In general terms, our research illustrates the need for and utility of controlled experimental designs of understudied variable elements in the grammar of BP.

References

- Lamberti, Luana, Kendra V. Dickinson & Eleni Christodoulelis. 2017. Variable Direct Objects in Brazilian Portuguese: Linguistic Constraints and Social Acceptability. Poster presented at NWAV 46, University of Wisconsin-Madison.
- Perini, Mário. 2010. *Gramática do português brasileiro*. São Paulo: Parábola.
- Schwenter, Scott A. 2014. Two kinds of Differential Object Marking in Portuguese and Spanish. In Patrícia Amaral & Ana Carvalho (Eds.), *Portuguese/Spanish interfaces. Diachrony, synchrony, and contact* (Issues in Hispanic and Lusophone Linguistics 1), 237-260. Amsterdam: John Benjamins.
- Zampaulo, André. 2014. Los pronombres de objeto directo *te* y *você*: variación lingüística en el portugués brasileño. *Cuadernos de Lingüística* 2, 173-193.